European Research Studies
Volume VI, Issue (3-4), 2003, pp. 143-162

Evaluation of Equity Mutual Funds' Performance Using a Multicriteria Methodology

K. Pendaraki[1], C. Zopounidis[2],

Abstract

The evaluation of the performance of mutual funds has been a very interesting research topic for not only researchers, but also for managers of financial, banking and investment institutions. In this study a well-known MCDA method based on the theory of outranking relations, the PROMETHEE II method (Preference Ranking Organisation Method for Enrichment Evaluations; Brans & Vincke, 1985) is used to develop outranking models for mutual funds' performance. This method is applied on real-world data of mutual funds derived from the Association of Greek Institutional Investors. The results of the PROMETHEE II method are indicative of ranking the funds from the best to the worst ones according to their performance.

Keywords: Mutual funds, Multicriteria analysis, Ranking.
JEL Classification: C52, E22, O16

1. Introduction

The mutual funds industry has experienced huge growth internationally in recent years, becoming the primary vehicle through which individuals and most institutions invest in capital markets. Within the European Association at present 21.576 mutual funds operate, with total assets rising to Euro 3.274 bn (data as of 31/03/2000, Association of Greek Institutional Investors). In the same way, the industry of collective investments in Greece is growing rapidly. According to recent data of the Association of Greek Institutional Investors (30/09/2000), there are 28 Mutual Fund Management Companies which are managing 248 mutual funds, with assets rising to GRD 11,48 trillion. A decade earlier (in 1990s), there were operating only 7 Mutual Fund Management Companies which were managing only 7 mutual funds with assets rising to GRD 147 billion.

[1] Konstantina Pendaraki, MSc., PhD Candidate, Technical University of Crete, Dept. of Production Engineering and Management, Financial Engineering Laboratory (http://www.dpem.tuc.gr.fel/), University Campus, 73100 Chania, Greece, Phone: +30-821-37366, Fax: +30-821-69410, E-mail: dina@ergasya.tuc.gr

[2] Professor Constantin Zopounidis, Technical University of Crete, Dept. of Production Engineering and Management, Financial Engineering Laboratory (http://www.dpem.tuc.gr.fel/), University Campus, 73100 Chania, Greece, Phone: +30-821-37236, 69551, Fax: +30-821-69410, 37236, E-mail: kostas@ergasya.tuc.gr, kostas@cha.forthnet.gr, Website: http//www.dpem.tuc.gr/fel/zopounidis.htm

This situation highlights the great growth of the Greek Mutual Fund Market. In contrast, in September of 2000 the number of the listed companies in the Athens Stock Exchange came up to 334 that are 34% more than the existed mutual funds. It is obvious that the variety of the attainable choices offered to every investor regarding the investments on mutual funds are more produced than the straightforward choices regarding the examination of the characteristics and the analysis of their performance. Indeed, this problem is more obvious in countries abroad. Suggestively, it is mentioned that the American Investment Company Institute counts more than 8.200 mutual funds when the listing companies in the Stock Exchanges of NYSE and NASDAQ of New York at the end of 1999 were about 7.800.

Thus, it is very difficult for investors to choose funds according to their decision policy, the risk levels that are willing to take, and their profitability needs. Today, in USA numerous business magazines, private firms, and financial institutions are specialized in giving regular rankings and ratings of mutual funds. Representative examples are the evaluations of funds given by Morningstar and the two well-known investors services of Moody's and Standard & Poor's, which greatly influence U.S. investor behaviour. According to Sharpe (1998), in USA the 90% of new money that are invested in stock funds in 1995 referred to funds that Morningstar gives four-star or five star ratings. Although this percentage may or may not be correct for mid-1998, certainly, there are few advertisements, which announce that a fund has received one star.

In Greece, there are no such institutions regarding the evaluation of mutual fund performance available to the Greek investors. The adoption of the evaluation systems of the foreign markets in Greek capital market is not feasible, such as these systems are based in specific category ratings that is possible not to be complied with the Greek market features. According to Sharpe (1998), such measures, like Morningstars, are appropriate measures to investors that place all their money in one fund. Morningstar makes the assumption that investors have some other basis for allocating funds and plan to use Morningstar's rankings in the case that they have to come up with a decision regarding which fund or funds to choose from each peer group. Thus, such measures are not appropriate performance measures when evaluating the desirability of a fund in a multifund portfolio, where the relevant measure of risk is the fund's contribution to the total risk of the portfolio.

The analysis of the nature and definition of risk in the portfolio selection and management shows that the risk is multidimensional and is affected by a series of financial and stock market data, qualitative criteria and macroeconomical factors which affect the process of the capital market. Many of the models used in the past are based on unidimensional approach that does not fit to the multidimensional nature of risk [Colson & Zeleny, (1979); Hurson & Zopounidis (1995)].

The empirical literature upon the evaluation measurements of the performance of mutual fund portfolios referred to Treynor index (1965), Sharpe's index (1966), Jensen's performance index (1968), Treynor-Mazuy model (1966), Henriksson-Metron model (1981), the CAPM, and several optimization models, etc. Eventhough these performance measurements adjusted to risk have been widely used in the assessment of portfolio performance, researchers have noted several restrictions

in their application, such as the use of a proxy variable of the theoretical market portfolio that can be criticized as inadequate, the evaluation of the performance of an investment manager for long and not short time periods, the acceptance of the assumption of borrowing and lending with the same interest rate, the validity of the Capital Asset Pricing Model, the consistency of the performance of investment managers over time, etc.

The multicriteria decision aid (MCDA) provides the requisite methodology framework in handling the problem of portfolio selection and management through a realistic and an integrated approach (Hurson & Zopounidis, 1997). MCDA methods incorporate the preferences of the decision-maker (financial/credit analysts, portfolio managers, managers of banks or firms, investors, etc.) into the analysis of financial decision problems. They are capable of handling qualitative criteria and are easily updated, taking into account the dynamic nature of the decision environment as well as the changing preferences of the decision-maker.

On the basis of the MCDA methodologies, this paper proposes the application of a method originated from the field of the multicriteria decision aid, the PROMETHEE II method (Brans & Vincke, 1985), in order to develop outranking models of the performance of mutual funds. The PROMETHEE II method is applied in a sample of Greek domestic equity mutual funds using data derived from the Association of Greek Institutional Investors for the period 1999-2000, in order to rank them from the best to the worst ones, or in other words to assign them within the highest and lowest performance positions.

The rest of the paper is organized as follows. Section 2 presents a brief literature review. Section 3 outlines the main features of the PROMETHEE II method. Section 4 is devoted to the application of PROMETHEE II method on our sample of equity mutual funds examining several different scenarios. Finally, section 5 concludes the paper and summarizes the main findings of this research.

2. Review of Past Empirical Studies

According to prior research, investors pay a great deal of attention to the selection of the mutual funds that will best accommodate their own financial situation (Morey M.R & Morey R.C, 1999). Thus, it is obvious that mutual funds classes are helping investors to choose funds according to their decision policy, the risk levels that are willing to take, and their profitability needs. Today, numerous business magazines, private firms, and financial institutions are specialized in giving regular rankings and ratings of mutual funds. Furthermore, there has been a wide variety of studies regarding the development of different models for the evaluation of the performance of mutual funds.

Friend, et al. (1962) have done the first extensive and systematic study of mutual funds. They created an index of five securities with the elements weighted by their representation in the mutual funds sample under consideration. According to their results, there is no strong relationship between turnover rates and performance. In 1966, there were written two papers that dominated in the area of mutual funds investment performance for the next twenty-five years. Sharpe (1966) in his

study calculated the reward-to-volatility ratio and found that the better performing funds tended to be those with the lower expenses. Furthermore, he showed that performance could be evaluated with a simple theoretically meaningful measure that considers both average return and risk. These results were very soon confirmed by the results of Jensen's research work (1968). He used the capital market line in order to calculate a performance measure (Jensen's alpha) for his data. Using this measure he concluded that the examined mutual funds were on average not able to predict security prices well enough to outperform the "buy-the market-and-hold" policy.

Lehmann and Modest (1987) in their research work tried to ascertain whether conventional measures of abnormal mutual fund performance are sensitive to the benchmark chosen to measure normal performance. They employed the standard Capital Asset Pricing Model (CAPM) benchmarks and a variety of Arbitrage Pricing Theory (APT) benchmarks in order to give an answer to the previous question. Cumby and Clen (1990) examined the performance of internationally diversified mutual funds. They used two performance measures, the Jensen measure and the positive weighting measure, proposed by Grinblatt and Titman (1989), and found that there is no evidence that funds provide investors with performance that surpasses that of a broad, international equity index over the examined period.

Brockett et al. (1992) in their empirical analyses of mutual fund investment strategies used a chance constrained programming approach in order to maximize the possibility of the performance of a mutual fund portfolio to exceed the performance of the S&P 500 index formalizing risk and return relations. Grinblatt and Titman (1994) examined the sensitivity of performance inferences to benchmark choice, they compared the Jensen measure with two new measures that were developed in order to overcome the timing-related biases of the Jensen measure, and finally they analyzed whether mutual fund performance is related to fund attributes. They concluded that the measures generally yield similar inferences when using different benchmarks and the tests of fund performance that employ fund characteristics suggest that turnover is significantly positively related to the ability of fund managers to earn abnormal returns.

Chiang et al. (1996) used an artificial neural network method in order to develop forecasting models for the prediction of end-of-year net asset values of mutual funds, taking into account historical economic information. They compared their forecasting results to those of traditional econometric techniques and concluded that neural networks significantly outperform regression models in situations with limited data availability. Murthi et al. (1997) examined the efficiency of mutual fund industry by different investment objectives. They tried to overcome the limitations of traditional indices, proposing a new measure of performance that is calculated through the data envelopment analysis. O' Neal (1997) in his research work tried to investigate whether the investors can receive diversification benefits from holding more than a single mutual fund in their portfolios. The results given by the simulation analysis that he conducted showed that the time-series diversification benefits are minimal but that the expected dispersion in terminal-period wealth can be substantially reduced by holding multiple funds.

Into their study, Indro et al. (1999) used artificial neural networks in order to predict mutual fund performance. Precisely, they used the fund's five-year annualized return, the turnover of the fund's portfolio, the price-earnings ratio, the price-book ratio, the median market capitalization, the percentage of cash and the percentage of stock (these ratios in relation with fund's portfolio) to predict the mutual fund performance, which is measured by the fund's risk-adjusted return. They used a multi-layer model and a nonlinear optimizer taking into account fund-specific historical operating characteristics in order to forecast mutual funds' risk adjusted return. They concluded that whether the neural network approach is superior to linear models for predicting mutual fund performance depends on the style of the fund. Morrey and Morrey (1999) in their empirical analysis used two basic quadratic programming approaches in order to identify those funds that are strictly dominated, regardless of the weightings on the different time horizons examined, relative to their mean returns and risks. Furthermore, these approaches endogenously determine a custom-tailored benchmark portfolio to which each mutual fund's performance is compared.

Cromwell et al. (2000) examined the change in the first three moments of the return distribution (mean, standard deviation, and skewness) of a portfolio of mutual funds as the number of funds in the portfolio increases. They concluded that diversifying across mutual funds substantially reduces portfolio dispersion but also causes undesirable increase in negative return skewness. Dalhquist et al. (2000) studied the relation between fund performance and fund attributes in the Swedish market. They examined 130 equity mutual funds for the period 1993-97. According to their work, performance is measured as the alpha in a linear regression of fund returns on several benchmark assets, allowing for time-varying betas. They came up with the conclusion that good performance occurs among small equity funds, low fee funds, funds whose trading activity is high and in few cases funds with good past performance. Wermers (2000) in his study performed a comprehensive analysis of mutual fund industry through a new database that allows an analysis of mutual funds in both the stock holdings level and the net return level from 1975 to 1994. He decomposed performance into several components to analyze the value of active fund managers. According to the results of the application of the performance decomposition methodology (characteristic selectivity and timing measures, average style measure, and execution costs) followed in this study, funds that hold stocks outperform the market, whereas their net returns underperform the market. Thus, funds include stocks to cover their costs. Finally, there is evidence that supports the value of active mutual fund management.

Ahmed (2001) evaluated various models in order to determine their efficiency in forecasting correlation among 202 equity mutual funds over the period 1979 to 1999. Correlation among funds is a very important determinant of portfolio risk. Precisely, this study forecasts mutual fund correlation using eight models (historical, mean and index models). According to their results, a Multi-Style Index, a Dynamic model and the Fama-French 3-Factor model presented the lowest prediction errors. Furthermore, the relative ranks of Multi-Style Index and Fama-French 3-Factor models have lower dispersion across different forecasting time period and in sub-

samples of funds belonging to similar or different style categories. The inference of these results is important for managing the performance of a portfolio of mutual funds. Gruber (2001) in his study identified the risk structure of mutual fund returns for 270 funds over the period 1985-1994 (in factor analysis) and for 274 funds over the period 1985-1995 (in cluster analysis). Precisely, he used a four-index model employing the S&P Index, and publicly available size, growth and bond indexes in order to examine what influences generate mutual fund returns and develop a model for measuring performance. He used factor analysis and proved that a fifth index appears to be present. In the case where he tested a publicly available index of growth mutual fund performance he found out that it explains a large proportion of the residuals from a four-index model. Finally, the data suggested that cluster analysis could be best used as an added influence to the based model. On the other hand, adding an index based on the dividend yield value index to the base model with a Morningstar Growth Fund Index explains correlation in a better way.

Zopounidis and Pendaraki (2002) into their study presented an integrated multicriteria decision aid methodology for the portfolio selection and composition problem in the case of equity mutual funds over the period 1997-1999. The methodology used consists of two stages. In the first stage the mutual funds are ranked according to their performance through PROMETHEE II method (originated from the field of the outranking relations of MCDA) based on several different weighting scenarios, in order to construct a portfolio consisting of a limited set of the best funds. In the second stage of this methodology it was applied a continuous MCDA technique through a goal programming formulation in order to solve the mutual funds portfolio composition problem specifying the proportion of each fund in the constructed portfolio. The proposed integrated approach constitute a significant tool that can be used to provide answers to two vital questions: (a) which funds are the most suitable to invest, and (b) what portion of the available capital should be invested in each one of these funds.

3. The PROMETHEE II Method

A decision-maker that is solving a multicriteria problem can take into account three kinds of methods: the aggregation methods using utility functions, the interactive methods and the outranking methods. In this paper is applied an outranking method called PROMETHEE based on the theory of outranking relations. The outranking methods include two phases: (a) the construction of an outranking relation, and (b) the exploitation of this relation in order to assist the decision-maker.

Brans (1982) was firstly proposed the PROMETHEE method. His study was followed by the studies of Brans and Vincke in 1985 and Brans, Vincke and Mareschal in 1986, which are among the most important publications of this method. PROMETHEE is a very simple method and easily understood by the decision-maker. Through this method, a valued outranking relation can easily be built, based on the extensions of the notion of criterion. The extended criteria used represent the natural notion of intensity of preference and the parameters that have to be fixed have real economic meaning. The basic principles of the PROMETHEE method

in relation with other methods of the same field are the following: (1) extension of the notion of criteria, (2) valued outranking relation, (3) exploitation of the outranking relation.

As far as the principle of the extension of the notion of criteria concerns, new criteria functions are proposed to the decision-maker such as, the usual criterion, the quasi-criterion, the criterion with linear preference, etc. In PROMETHEE method the valued outranking relation is less sensitive to small modifications and its interpretation is easy. The exploitation of the valued outranking relation of the PROMETHEE method refers to the case in which the alternatives (mutual funds) have to be ranked from the best to weakest (best performance to worst performance). There are two PROMETHEE methods: (1) the PROMETHEE I method that provides a partial preorder on the set of possible alternatives, and (2) PROMETHEE II that provides a total preorder on the set of possible alternatives. As it is already mentioned above, in this case study the PROMETHEE II method is applied in order to develop outranking models of the performance of mutual funds.

Brans et al. (1986) used six types of functions that cover most of the cases occurring in practical applications. For any two alternatives α and b, the function H(d) (figure 3.1) is specified as follows:

$$H(d) = \begin{cases} P(a,b), d \geq 0 \\ P(b,a), d \leq 0 \end{cases}$$

where $P(a,b)$ is the preference function of mutual fund a with regard to mutual fund b.

The six types of the preference function $P(a,b)$ (generalized criterion) which are used for the determination of function $H(d)$ are presented in table 3.1. For each generalized criterion, only a few parameters (maximum 2) have to be identified by the decision-maker.

For each couple of mutual funds a and b, is defined a preference index of mutual fund a with regard to mutual fund b over all criteria, when all criteria have the same importance. This preference index determines the valued outranking relation and is defined as follows:

$$\Pi(a,b) = \frac{\sum_{i=1}^{k} \pi_i P_i(a,b)}{\sum_{i=1}^{k} \pi_i}$$

where π_i is the weight of each criteria g_i ($i = 1, ..., k$).

The preference index varies from 0 to 1 and defines a complete valued relation that presents the global intensity of preference between couple of mutual funds (alternatives). When $\Pi(a, b) \approx 0$ we have weak preference of mutual fund a with regard to mutual fund b for all criteria. On the other hand when $\Pi(a, b) \approx 1$ we have strong preference of mutual fund a with regard to mutual fund b for all criteria.

Regarding the exploitation of the outranking relations for the ranking of mutual funds (alternatives) the following two preference flows are defined:

- the outgoing flow ➲ $\varphi^+(a) = \sum_{b \in k} \Pi(a,b)$, (measures how much mutual fund a is preferred to all other funds), and
- the incoming flow ➲ $\varphi^-(a) = \sum_{b \in k} \Pi(b,a)$, (measures how much the other mutual funds are preferred to mutual fund a),

where k is the total number of mutual funds.

Let now consider for each mutual fund α the following net-flow:

$$\varphi(a) = \varphi^+(a) - \varphi^-(a)$$

In PROMETHEE II method this net-flow is used for ranking the funds:

- aPb [a outranks b if $\varphi(a) > \varphi(b)$],
- aIb [a is indifferent to b if $\varphi(a) = \varphi(b)$]

This is the PROMETHEE II complete relation through which all mutual funds of K are completely ranked. Finally, it must be mentioned that the valued outranking relation is presented through a valued outranking graph that gives the ranking of all the examined mutual funds.

4. Application

4.1. Data Set Description and Criteria

The data set used to examine the performance rankings of mutual funds (alternatives) consists of monthly data of all domestic equity mutual funds over the period 1999-2000. This data set is derived from the Association of Greek Institutional Investors and refers to the monthly net asset value and the monthly return of mutual funds. Further information was derived from the Athens Stock Exchange and the Bank of Greece, regarding the monthly return of market portfolio and the monthly return of Treasury bill respectively. This case study includes those mutual funds that operate the whole monthly period for each year examined. The rest mutual funds were excluded from the analysis because they had not been in existence long enough in order their performance to be meaningful for the period examined.

This restriction is imposed in order to have complete records. In this way the sample for 1999 data is reduced from 53 funds that are at the end of December 1999 in only 34 that pass our restriction, but it should be noted that our sample includes 81.67% of the shares in the funds classified as "Growth" funds. The other 19 funds are introduced during the year and there is no full record for them. Furthermore, we traced out the changes of the names in the funds in order to have a better presentation of the category. At the end of the year 2000 there are 78 funds but only 51 funds pass our criteria. The funds in our sample represent 95.03% of the shares in this category. Again we traced out the changes of the names and splits of the funds. The sample is affected, even though in a very limited way, by

survivorship bias. In fact, even if the data set captures the changes in the name of the mutual funds, the selection procedure does not include the funds that changed investment policy.

The criteria that were used to evaluate mutual fund performance in annual base are: (1) annual percentage change of net asset value of mutual fund (from the last month of the previous year), (2) beta (β) coefficient, (3) value at risk, (4) annual return, (5) Treynor index (1965), (6) Sharpe index (1966), (7) Jensen's (1968) alpha (a) coefficient.

The beta (β) coefficient is a measure of fund risk in relation to the market risk. It is called systematic risk and the asset pricing model implies that is crucial in determining the prices of risky assets. The beta coefficient is defined as follows:

$$\beta = \text{cov}(R_{pt}, R_{Mt})/\text{var}(R_{Mt})$$

where

$\text{cov}(R_{pt}, R_{Mt})$ = covariance of monthly return of mutual fund with market portfolio

$\text{var}(R_{Mt})$ = variance of monthly return of market portfolio.

If $\beta > 1$ we have an aggressive fund which gives larger return in an increasing market but greater losses in a total diminish of the price level. If $\beta < 1$ we have an defensive fund which gives lower risk and its returns are changing more smoothly than market changes.

Another well-known measure of risk is the Value at Risk. Value at risk is also known as VAR, and its popularity was much enhanced by the 1993 study by the Group of Thirty, Derivatives: Practices and Principles, which strongly recommended VAR analysis for derivatives trading. The VAR measure gives an answer in the question " How much can the value of a portfolio decline with given probability in a given time period?". The calculation of VAR is based on certain assumptions about the statistical distribution of the fund's return. Precisely, in order VAR to be calculated the assumption that returns follow normal distribution is done. One of the properties of normal distribution is that 95% of all observations occur within 1.96 standard deviations meaning that the probability of an observation to fall outside the 1.96 standard deviations below the mean is only 5%. In the case of VAR calculation, only losses are taken into account and the relevant probability is only 2.5%. The VAR measure is defined as follows: Monthly VAR = Mean Monthly Return - 1.96 Standard Deviation of Monthly Return. The power of VAR models refer to the opportunity that give for the construction of a measure of risk for a portfolio not from its own past volatility but from the volatilities of risk factors affecting the portfolio as it is constituted today. It is a measure highly correlated with volatility because it is proportional to standard deviation.

The traditional total performance measures, Sharpe index, and Treynor index are used to measure the expected return of a fund per unit of risk. These measures are defined as follows:

$$\text{Sharpe index} = (R_p - R_f)/\sigma_p$$

$$\text{Treynor index} = (R_p - R_f)/\beta_p$$

where

R_p = annual return of mutual fund
R_f = annual return of Treasury bill (annual risk free interest rate)
σ_p = standard deviation of monthly return of mutual fund (total risk of mutual fund)
β_p = systematic risk of mutual fund.

The Sharpe index or alternatively the reward-to-variability ratio is a useful measure of performance. It is defined as the difference between the fund's annual return and the pure interest rate to the standard deviation of the annual rate of return. In other words, the numerator shows the reward provided by the investor for bearing risk, while the denominator shows the amount of risk actually bear. It is obvious that this ratio is the reward per unit of variability. Furthermore, Sharpe index represents a relevant measure of mutual fund performance for investors who are not well diversified and, therefore, are concerned with their total risk exposure when evaluating mutual fund performance. The Sharpe performance measure reflects both the differences in returns to each fund and the level of mutual fund diversification.

The Treynor index is obtained by simply substituting variability (the change in the rate of return on a fund associated with 1 per cent change in the rate of return on, say, the market portfolio) by volatility in the formula of the Sharpe index. Thus, the Treynor index is similar to the Sharpe index except that performance is measured as the risk premium per unit of systematic (β_p) and not of total risk (σ_p). The evaluation of mutual funds with those two indices show that a mutual fund with higher performance per unit of risk is the best managed fund, while a mutual fund with lower performance per unit of risk is the worst managed fund.

The Jensen alpha measure is the intercept in a regression of the time series of fund excess returns against the time series of excess returns on the benchmark. Both the Treynor index and the Jensen alpha assume that investors are well diversified and, therefore, they are only taking into account systematic risk when evaluating fund performance. The Jensen alpha measure is given by the regression of the following model:

$$\text{Jensen model: } (R_{pt} - R_{ft}) = \alpha_p + \beta_p (R_{Mt} - R_{ft}) + \varepsilon_p$$

where

R_{pt} = monthly return of mutual fund
R_{ft} = monthly return of Treasury bill (monthly risk free interest rate)
R_{Mt} = monthly return of portfolio market
α_p = Jensen alpha measure
β_p = estimated risk parameter
ε_p = error term (independent normally distributed random variable with $E(\varepsilon_p) = 0$).

The coefficient α_p will be positive if the manager has any forecasting ability and zero if he has no forecasting ability. On the other hand, we can rule out a negative coefficient α_p by perversing forecasting ability.

4.2. Presentation of Results

Our target in mutual funds' performance analysis is to rank the examined mutual funds using the PROMETHEE II method for two years separately. Generalization of the results was quite difficult because the economic environment, which prevailed in the year 2000, was completely different to the one that mutual funds enjoyed in 1999. Most of the mutual funds in our samples managed to realize higher returns than those of the market during the year 1999, whereas during the year 2000 only few funds were successful.

In general, we found that few mutual fund managers in 1999 earned some "abnormal" returns as measured by Jensen alpha, but this was not the case in 2000, when we found negative "abnormal" returns. Generally speaking, we found no evidence of manager's ability to earn "abnormal returns" with exception of few funds for which there was some evidence of "abnormal" returns, either positive or negative.

It was revealed that while the mutual funds in our samples had higher returns for unit of variability than those of the market, as measured by the Sharpe ratio for the year 1999, in 2000 they were not able to repeat the same performance. The same result is valid for returns per unit of volatility, as measured by Treynor index. The funds in our sample were found to be more risky than the market in 1999, in terms of risk as measured by the beta coefficient and the value at risk, whereas in 2000 there were fund to be less risky than the market. In general, we found out that mutual funds were less than perfectly diversified and were not able to fully take advantage of the benefits that come from diversification.

PROMETHEE II method gives a complete rank of the mutual funds examined in the two years period. Three crucial issues in using this method are the selection of the generalized criteria, the specification of the criteria's weights, and finally the evaluation of the parameters of each generalized criterion. The results obtained from the application of this method show that a mutual fund with the highest average ranking is the best one, while a mutual fund with the lowest average ranking is the worst one.

Precisely, in this case study we examined the following three generalized criteria: (a) the Gaussian criterion, (b) the criterion with linear preference, and (c) the criterion with linear preference and indifference area. We believe that these three generalized criteria cover most of the cases occurring in practical applications and the majority of the behaviors of the decision-maker. Tables 4.2.1 and 4.2.2 represent the average rankings of the performance of the examined mutual funds for 1999 and 2000 respectively, obtained through PROMETHEE method according to the three aforementioned generalized criteria.

For all the generalized criteria taking into account in this case study, 50 random weight combinations for each of the seven performance measures were used. These

random weight combinations were used taking also into account the lack of an expert stock market analyst that could determine the weights of the performance measures used in this study. Each weight is a random variable that is uniformly distributed into the interval [1, 100]. It is very important to minimize the randomness of the weights, so we assigned the weights of the criteria used according to their significance and importance. Precisely, we assigned weights in the way that five criteria will equally count for 15 % in the final ranking with exception of the alpha coefficient and the annual percentage change of net asset value that will count for 12,5% each. We decided to give less weight to these two criteria because of the alphas insignificance in most of the cases and the high variability that the percentage change of asset value is presented in the examined period.

For each weight combination the parameters σ, p and q were defined. In the Gaussian criterion, parameter σ was defined as $\sigma = 0.25 \times s \times n$ *(step=0.25), where s = standard deviation of each criterion, and n = 1,2,...,10. According to this generalized criterion 500 different scenarios were examined. In criterion with linear preference, parameter p was defined as* $p = 0.25 \times s \times n$ *(step=0.25). Once again, 500 different scenarios were examined. Finally, in the criterion with linear preference and indifference area, the parameter q was defined as* $q = 0.15 \times s \times n$ *(step=0.15), and the parameter p was defined as* $p = q + 0.25 \times s \times n$ *the (step = 0.15). According to this generalized criterion 6250 different scenarios were examined.*

The average rankings for each year were obtained over the rankings of the examined year for all the different scenarios used into the analysis and compared with the Kendall's coefficient of concordance (Kendall's W). Kendall's W is a statistical measure which examines the agreement over the rankings which are obtained according to the generalized criterion that is used in each time. The Kendall's coefficient of concordance (Siegel, 1956) is given by the following ratio:

$$W = \frac{\sum_{j=1}^{n}(R_j - \frac{1}{2}k(n+1))^2}{\frac{1}{12}k^2(n^3 - n) - k\sum T}$$

where k are the scenarios performed for the evaluation of n mutual funds and for each of the n mutual funds the k ranks are totalled, giving rank sums R_j, for $j=1,2,...,n$ and $T = \sum(t^3 - t)/12$, each t being the number of occurrences of each tied rank within a scenario, and the summation of T being over all scenarios containing ties. The values of this coefficient fall in the interval between 0 and 1. If the coefficient takes the value 1 this means that the rankings are the same, or in other words that there is a perfect agreement between all the rankings. The results of the Kendall's W coefficient and its Chi-square value are presented in the final rows of the two tables.

According to the results obtained there are no significant differences in the rankings of the mutual funds for the examined years, because the Kendall's W coefficient takes high values in all the cases examined in this research. The highest stability in the rankings is met in the results obtained from the Gaussian criterion

(0.98, 0.99 in 1999 and 2000 respectively), while the lowest in the results obtained from the criterion with linear preference (0.66 and 0.73 for 1999 and 2000 respectively). Chi-Square is very high and fully supports the significance of the rankings in all the cases for both years. These rankings show the position of a given mutual fund in terms of performance compared with the other mutual funds.

The final rankings output on the three generalised criteria used in PROMETHEE II method were the same regarding the best performers and the worst performers. We noticed only some changes in the mid-class performers. The rankings were different for each year but some mutual funds gave signals of consistence in their performance in the sense that the funds, which were the best in the year 1999, had also relatively good performances in the year 2000. The same applied to the worst performers. Also, it was observed that most of the new entrants in our sample for the year 2000 were among the worst performers.

Precisely, according to the three generalized criteria examined, in 1999 there are four mutual funds that are ranked in all cases within the four highest positions (table 4.2.1). These mutual funds are: Eurohellenic Equities, Dorian, Hermes Dynamic and ATE Growth. On the other hand, for the same year the mutual funds that are ranked in all cases in the two lowest positions are: Cretafund Growth and International.

In 2000, Barclays FTSE/ASE 20 is ranked in all cases (table 4.2.2) within the first highest position. According to the Gaussian criterion, the CityFund Equity is ranked in the second highest position while according the other two generalized criteria the Creta fund is ranked in the second highest position. International, General Small Cap., and International Index Midcap FTSE/ASE MID-40 are ranked in all cases in the three lowest positions.

5. Conclusions and Discussion

Although the topic of mutual fund evaluation has been ignored for many years, it recently received considerably attention. Today, numerous business magazines, private firms, financial, banking and investment institutions are specialized in giving regular rankings and ratings of mutual funds. Furthermore, there has been a wide variety of studies regarding the development of different methods for the evaluation of the performance of mutual funds.

The aim of this study was to solve the ranking problem of the performance of mutual funds through PROMETHEE II method, originated from the field of the multicriteria decision aid. This method was applied to rank mutual funds from the best to the worst ones based on three generalized criteria and seven criteria of mutual fund performance through several different scenarios.

Kendall's W coefficient was used to examine the stability of the rankings obtained from the different weighting scenarios for each one of the two years examined. According to the results there are not big differences in the rankings of the mutual funds for the two examined years.

The results of this research work will be a guide in the decision aid process for every potential investor/manager of mutual funds and investor advisor. In particular,

the proposed method used in the development of the evaluation models of the performance of mutual funds gives a series of advantages in practical level. The flexibility of multicriteria methods and their ability to adapt in their developed models, the preferences, the experience and the policy of an investor/portfolio manager into the decision making process gives the opportunity to managers of financial, banking and investment institutions to study the problem of evaluating the performance of mutual funds and ranking funds from the best to the worst ones through an integrated approach. (e.g. the results of the rankings of the funds can be used for the construction of a portfolio consisted from the best funds showing to the investor which funds are the most suitable to invest and what portion of the available capital should be invested in each of these funds).

Further work in the development of evaluation models of the performance of mutual funds of all the categories (value, income, bond, international, etc.) through other MCDA methods such as the UTA method (UTilitis Additives; Jacquet-Lagrθze and Siskos, 1982), the UTADIS method (UTilitis Additives DIScriminantes; Jacquet-Lagrθze and Siskos, 1982; Zopounidis and Doumpos, 1999), the ELECTRE TRI method (Elimination Et Choix Traduisant la Realitι; Yu, 1992), and the AHP method (Analytic Hierarchy Process; Saaty, 1980) etc., is also very interesting to be conducted.

References

Ahmed, P., 2001, "Forecasting correlation among equity mutual funds", *Journal of Banking and Finance*, 25, 1187-1208.

American Investment Company Institute in http://www.ici.org.

Association of Greek Institutional Investors in http://www.ugii.gr.

Brans, J.P., 1982, "L' ingιnierie de la dιcision, l' ιlaboration d'instruments d' aide á la dιcision", Colloque d' aide á la dιcision (Universitι Laval, Quebec).

Brans, J.P. and Ph. Vincke, 1985, "Preference ranking organization method: The PROMETHEE method for multiple criteria decision making", *Management Science*, 31(6), 647-656.

Brans, J. P., Vincke, Ph and B. Mareschal, 1986, "How to select and how to rank projects: The PROMETHEE method", *European Journal of Operational Research,* 24, 228-238.

Brockett, P.L., Charnes, A. and W.W. Cooper, 1992, "Chance constrained programming approach to empirical analyses of mutual fund investment strategies", *Decision Sciences*, 23, 385-403.

Chiang, W.C., Urban, T.L. and G.W. Baldridge, 1996, "A neural network approach to mutual fund net asset value forecasting", Omega, Int. J. Mgmt Sci., 24, 205-215.

Colson, G. and M. Zeleny, 1979, "Uncertain prospects ranking and portfolio analysis under the condition of partial information" in Verlag Anton Hain [ed.]: *Mathematical Systems in Economics* 44, (Maisenheim).

Cromwell, N.O. Taylor, W.R.L. and J.A. Yoder, 2000, "Diversification across mutual funds in a three-moment world", *Applied Economics Letters*, 7, 243-245.

Cumby R.E. and J.D. Glen, 1990, "Evaluating the performance of international mutual funds", *The Journal of Finance*, XLV, 497-521.

Dahlquist, M., Engstrom, S. and P. Soderlind, 2000, "Performance and characteristics of Swedish mutual funds", *Journal of Financial and Quantitative* Analysis, 35 (3), (September), 409-423.

Friend, I., Brown, F., Herman, E. and D. Vickers, 1962, "A study of Mutual Funds", U.S. Securities and Exchange Commission.

Grinblatt, M. and S. Titman, 1989, "Portfolio performance evaluation: Old issues and new insights", *Review of Financial Studies,* 2, 393-421.

Grinblatt, M. and S. Titman, 1994, "A study of monthly fund returns and performance evaluation techniques", *Journal of Financial and Quantitative Analysis,* 29, 419-443.

Group of Thirty, 1993, "Derivatives: Practices and Principles", (Washington, DC).

Gruber, M. J. 2001, "Identifying the risk structure of mutual fund returns", *European Financial Management,* 7 (2), 147-159.

Henriksson, R and R. Merton, 1981, "On market timing and investment performance", *Journal of Business*, 54(4), 513-534.

Hurson, Ch. and C. Zopounidis, 1995, "On the use of Multicriteria decision aid methods for portfolio selection", *Journal of Euro-Asian Management,* 1/2, 69-94.

Hurson, Ch. and C. Zopounidis, 1997, "Gestion de portfeuille et Analyse Multicritθre", *Economica*, (Paris).

Indro, D.C., Jiang, C.X., Patuwo, B.E. and G.P. Zhang, 1999, "Predicting mutual fund performance using artificial neural networks", *Omega*, 27, 373-380.

Jacquet-Lagrθze, E. and Y. Siskos, 1982, "Assessing a set of additive utility functions for multicriteria decision making: The UTA method", *European Journal of Operational Research*, 10, 151-164.

Jensen, C.M., 1968, "The Performance of Mutual Funds in the Period 1945-1964", *Journal of Finance,* 23, 389-416.

Lehmann, B.N. and D.M. Modest, 1987, "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons", *The Journal of Finance,* XLII, (2), June, 233-265.

Morningstar in http://www.morningstar.com.

Moody's investor service in http://www.moodys.com.

Morey M.R. and R.C. Morey, 1999, "Mutual fund performance appraisals: A multi-horizon perspective with endogenous benchmarking", *Omega*, Int. J. Mgmt Sci., 27, 241-258.

Murthi, B.P.S. Choi, Y.K. and P. Desai, 1997, "Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach", *European Journal of Operational Research,* 98, 408-418.

O' Neal, E.S., 1997, "How many mutual funds constitute a diversified mutual fund portfolio", *Financial Analysts Journal*, (March/April), 37-46.

Saaty, T.L., 1980, "The analytical hierarchy process" in McGraw-Hill [ed.], (New York).

Sharpe, W.F., 1966, "Mutual Fund Performance", *Journal of Business,* 39, 119-138.

Sharpe, W.F., 1998, "Morningstar's risk adjusted ratings", *Financial Analysts Journal,* July/August, 21-23.

Siegel, S., 1956, "Nonparametric Statistics for the Behavioral Sciences", in McGraw-Hill [ed.], (New York).

Standard & Poor's investor service in http://www.moodys.com.

Treynor, J.L., 1965, "How to rate management of investment funds", *Harvard Business Review,* 43, 63-75.

Treynor, J.L. and K.K. Mazuy, 1966, "Can mutual funds outguess the market", *Harvard Business Review,* 131-136.

Yu, W., 1992, "ELECTRE TRI – Aspects Méthodologiques et Manuel d'Utilisation", *LAMSADE Laboratory,* document No 74, Université de Paris (Dauphine).

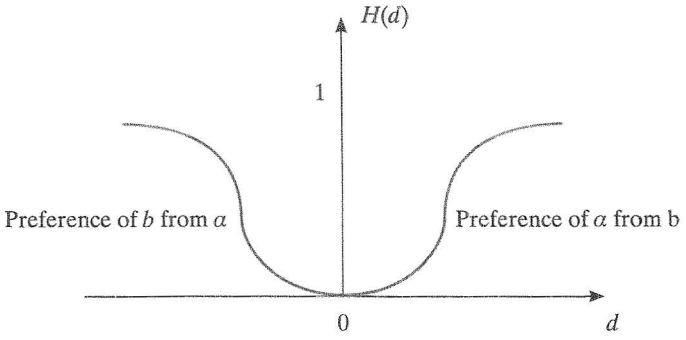
Wermers, R., 2000, "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses", *The Journal of Finance,* LV, (4), (August), 1655-1703.

Zopounidis, C. and M. Doumpos, (1999), " A multicriteria decision aid methodology for sorting decision problems: The case of financial distress", *Computational Economics,* 14/3, 197-218.

Zopounidis, C. and K. Pendaraki, 2002, "An integrated approach on the evaluation of equity mutual funds' performance", *European Journal of Business and Economic Management,* (in press).

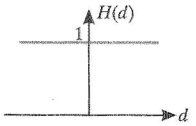
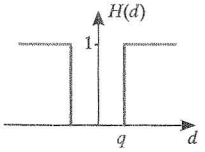
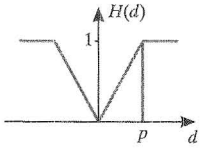
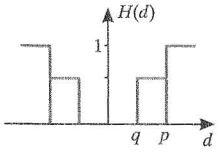
FIGURES

Figure 3.1
H(d) function



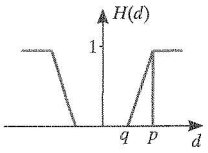
TABLES

Table 3.1: *Six types of generalized criteria*

Type of generalized criteria	Preference function H(d)=P(a, b)	Parameters
I. Usual criterion $H(d)=\begin{cases}0 \text{ if } d=0\\1 \text{ if } d\neq 0\end{cases}$		-
II. Quasi-criterion $H(d)=\begin{cases}0 \text{ if } -q\le d\le q\\1 \text{ if } d<-q \text{ or } d>q\end{cases}$		q
III. Criterion with linear preference $H(d)=\begin{cases}0 \text{ if } -p\le d\le p\\1 \text{ if } d<-p \text{ or } d>p\end{cases}$		p
IV. Level Criterion $H(d)=\begin{cases}0\\1/2\\1\end{cases}$		q, p

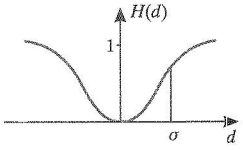
V. Criterion with linear preference & indifference area

$$H(d)=\begin{cases}0 & \text{if } |d|\le q\\ (|d|-q)/(p-q) & \text{if } q<|d|\le p\\ 1 & \text{if } p<|d|\end{cases}$$



q, p

VI. Gaussian criterion
$H(d) = 1 - \exp\{-d^2/2\sigma^2\}$



σ

Table 4.2.1: *Average rankings of mutual funds in 1999*

MUTUAL FUNDS	Gaussian criterion	Criterion with linear preference	Criterion with linear preference and indifference area
Eurohellenic Equities	33.9	33	33.4
Dorian	33.1	32	32.3
Hermes Dynamic	32	29.7	29.1
ATE Growth	30.7	28.6	28.2
A. Trust New Enterprises	29.5	25.9	28.6
Midland Hellinobret.	29.3	22.7	25
Ioniki Growth	28.3	26.9	27.4
Olympia Growth	27.2	23.9	25.8
GENERAL Growth Dom.	25.6	23.7	22.9
Interamerican Devel.Co	24.8	23.4	26.5
Alpha Trust Growth	24.3	24.8	26.9
Delphi	23.1	22.5	21.3
Xios Growth	22.1	13.6	15.5
Interamerican Dynamic	20.8	17.1	18.1
Delos Blue Chips	20.3	21.7	19.7
Alpha Growth	19	23.2	21.2
Metrolife Growth	18	15.5	17
Nat. Nederlanden	16.6	10.4	12
Barclays FTSE/ASE 20	15.5	10.5	10.5
LAIKH TELESIS	15.4	11.3	13.3
Cyprus Greek	13.9	15.1	13.6

Sorgen Invest	12.8	7.4	7.4
Aspis Pronia	11.5	13.1	10.3
European Reliance	10.2	11.3	11.1
Allianz Growth	9.5	18.6	17.1
ABN AMRO Growth	8.2	13.1	12.1
Alpha Trust Infrastructure	7.8	16.8	16.4
Ioniki Athens Index Fund	7.4	7.2	8.4
Ergasias Growth	6.8	8.9	8.4
M/F ETBA P&K Capital	6.3	15.5	13.4
Alico Growth	5	12.4	10.9
Sigma Greek Index Fund	3.1	7.2	5.3
Cretafund Growth	2	5.4	4.4
International	1	2.7	1.5
Kendall Coeff. of Concordance	**0.98**	**0.66**	**0.74**
Chi-Square	*325.29*	*10993.16*	*129510.94*

Table 4.2.2: *Average rankings of mutual funds in 2000*

MUTUAL FUNDS	Gaussian criterion	Criterion with linear preference	Criterion with linear preference and indifference area
Barclays FTSE/ASE 20	50.9	48	49.2
CitiFund Equity	49.8	45.6	48.1
Creta Fund	49.2	47.9	48.9
Delos Blue Chips	47.7	45.7	45.4
Delphi	46.5	40.8	43
Egnatia Theseas FTSE ASE 20	45.9	46.5	47.2
Alpha Growth	45.8	45	45.1
Sogen Invest	44.2	36.4	39.3
Sigma Athens Index Fund	42.7	40.1	42
Egnatia Athena Growth	42.1	38	40.9
Interamerican Dynamic	41.2	40.6	40.7
ATE Growth	39.6	38.8	37.9
European Reliance	39.3	38.5	38.6
Nat. Nederlanden	38.1	29.3	30.8
Alpha Trust Infrastructure	36.6	34.1	34.8
Allianz Growth	36.1	36.4	35.6
Eurohellenic Equities	34.6	30.7	31.1

Ioniki Athens Index Fund	34	35.7	34.8
International Index FT/S&P	32.7	30	32.3
Ergasias Growth	32.6	30.2	32.8
Xios Growth	31.1	22.6	24.4
ABN AMRO Growth	30	32.3	30.9
Dorian	28.8	28.2	27.5
Olympia Growth	27.8	19.5	22.3
M/F ETBA P&K Capital	26.4	20.7	19
Hermes Dynamic	26.2	29	28.6
European Reliance Devel.Co	25.2	26.6	24.3
Metrolife Growth	23.7	22	21.7
Aspis Pronia	23.5	25.9	26.2
Piraeus- Proteus	21.8	20.2	19.2
Barclays Growth	21.3	23.9	23.2
Alico Growth	20.1	27.3	25.8
Ioniki Growth	18.9	21	20.4
A. Trust New Enterprises	18.6	23.6	23.1
Allianz Aggressive	16.8	24.9	23.9
XIOS Small Cap	15.8	10.4	12
International MF Real Estate	14.4	15.4	15.1
ERGO Dynamic	14.3	18.6	15.9
Interamerican Devel.Co	13.2	13.3	14.9
Delphi Small Cap	11.5	13.4	12
Alpha Trust Growth	11.3	18	16.2
Interamerican Olympionikis	10.3	15.3	12.7
NEXUS - International	8.7	6.6	7.8
Alpha Dynamic	7.3	12	11.5
Cyprus Greek	6.7	12.7	12.4
Nat. Nederl Dynamic Co.	6.5	11.5	8.2
LAIKH TELESIS	6.2	8.8	8.6
GENERAL Growth Dom.	4	10.7	10.4
International	3	4.7	4.1
GENERAL Small Cap.	2	5.2	2.4
International Index Midcap FTSE/ ASE MID-40	1	3.3	2.4
Kendall Coeff. of Concordance	**0.99**	**0.73**	**0.81**
Chi-Square	*497.87*	*18272.38*	*213554.40*